

**ON A COMPLETENESS PROBLEM FROM THE
PAPER “CONTRACTIVE PROBABILITY METRICS
AND ASYMPTOTIC BEHAVIOR OF DISSIPATIVE
KINETIC EQUATIONS” BY J. A. CARRILLO, G.
TOSCANI, RIV. MAT. UNIV. PARMA (7) 6 (2007),
75-198**

MALGORZATA STAWISKA

ABSTRACT. We prove completeness of the space $\mathcal{P}_s^=$ of probability measures in \mathbb{R}^N which have equal moments up to order s , where $s \in \mathbb{N}$, endowed with the metric $d_s(\mu, \nu) = \sup_{x \in \mathbb{R}^N \setminus 0} \frac{|\hat{\mu}(x) - \hat{\nu}(x)|}{|x|^s}$. This solves an open problem formulated in [CT].

MSC 2010: Primary 60B10; Secondary 60E10

Keywords: convergence of probability measures; Fourier transform; moments; completeness.

1. INTRODUCTION

The possibility of introducing a metric on probability measures over a Polish space which metrizes their weak*-convergence is a topic with a long history. Many examples of such metrics are known (for collections of such examples, see e.g. [Du02], [GSu], [V]), exploring properties of various quantities and objects associated to probability measures. The systematic use of metrics based on the Fourier transforms of measures (and more generally, of tempered distributions) was initiated in the 1950s by J. Deny ([De50], [De51]). Developing and generalizing ideas in potential theory due to H. Cartan ([Ca41], [Ca45]), he studied the metric given for two probability measures μ, ν in \mathbb{R}^N by the L^2 norm of the Riesz potential of their difference: $d(\mu, \nu) = \|\frac{\hat{\mu} - \hat{\nu}}{|\cdot|^{s/2}}\|_2$, with $0 < s < N$. Some of his arguments carry over to the case of $s = N$ (the logarithmic potential in \mathbb{R}^N) and were also made explicit in [CKL]. Recently a new class of metrics dependent on a real positive parameter $s > 0$, motivated by applications to statistical physics and optimal transport theory ([GTW], [TV], [CT]), was defined by using Fourier transforms of probability measures. The aim of our present note is to prove a theorem concerning completeness of a certain space

Date: February 28, 2017.

of probability measures endowed with this kind of metric. This was posed as an open problem in [CT]. All relevant definitions, auxiliary results, the statement of the theorem and finally its proof are presented in the next two sections.

We should mention that another class of Fourier-based probability metrics was lately introduced by Cho ([Cho15]). These metrics are also related to the (inverse) Riesz potentials of Fourier transforms of the measures, but the discussion is beyond the scope of this note.

2. PRELIMINARY MATERIAL

Throughout, we will work in \mathbb{R}^N , $N \geq 1$. For a function f which has derivatives up to order $m - 1$ in an open set $U \subset \mathbb{R}^N$ and the derivative of order m at $0 \in U$, we will use Taylor's formula with remainder in Peano form ([Sch81a], Theorem 37):

$$f(x) = f(0) + f'(0).x + \dots + \frac{f^{(m)}(0)}{m!}.(x, \dots, x) + \alpha \|x\|^m,$$

where α tends to 0 together with x .

The converse will be also useful:

Theorem 2.1. ([Sch81a], Theorem III.7.38): *Let f be a m -fold differentiable mapping from an open set $U \subset E$, where E is an affine normed space, into a normed vector space F . If there exist k -linear continuous symmetric mappings L_k of the space E^k into F , $k = 1, \dots, m$, and an element $L_0 \in F$ such that*

$$f(a + h) = L_0 + L_1 h + \frac{L_2}{2!} h^2 + \dots + \frac{L_m}{m!} h^m + \alpha \|h\|^m,$$

where $\alpha = \alpha(h)$ tends to 0 along with h , then necessarily

$$L_k = f^{(k)}(a), \quad k = 1, \dots, m$$

and

$$L_0 = f(a).$$

By a “probability measure on \mathbb{R}^N ” we mean a (Radon) probability measure on the Borel σ -algebra generated by the standard topology in \mathbb{R}^N . We will use some standard definitions and results in probability:

Definition 2.2. (cf. e.g. [Du02], Section 9) Let $\mathcal{C}^b(\mathbb{R}^N)$ be the set of all bounded continuous real-valued functions on \mathbb{R}^N . We say that the probability measures μ_n converge (weakly*) to a probability measure μ if and only if for every $\varphi \in \mathcal{C}^b(\mathbb{R}^N)$, $\int \varphi d\mu_n \rightarrow \int \varphi d\mu$ as $n \rightarrow \infty$.

Notation 2.3. The function

$$\hat{\mu}(x) = \int_{\mathbb{R}^N} e^{-i\langle x, v \rangle} d\mu(v), \quad x \in \mathbb{R}^N,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in \mathbb{R}^N , is the Fourier transform (characteristic function) of μ .

Note that the function $\hat{\mu}$ is continuous and bounded on \mathbb{R}^N . Furthermore (cf. [Du02], Theorem 9.4.4), if $\int_{\mathbb{R}^N} x^\beta d\mu(x)$ is finite for a multi-index $\beta = (\beta_1, \dots, \beta_N)$, then $\hat{\mu}$ has continuous partial derivative $D^\beta \hat{\mu}$ in \mathbb{R}^N , satisfying the formula $D^\beta \hat{\mu}(x) = \int_{\mathbb{R}^N} (iv)^\beta e^{-i\langle x, v \rangle} d\mu(v)$.

Recall also that the existence of moments $\int x^\beta d\mu$ up to order s of a probability measure μ implies the existence and continuity of partial derivatives $D^\beta \hat{\mu}$ of the Fourier transform $\hat{\mu}$ of μ for all $|\beta| = \beta_1 + \dots + \beta_N \leq s$. It follows that (cf. [Sch81a], Ch. III, Section 5, Theorem 28) $\hat{\mu}$ is s times continuously differentiable in \mathbb{R}^N .

An important relation between the weak* convergence of probability measures and their Fourier transforms is given by the Lévy continuity theorem:

Theorem 2.4. (cf. [Du02], Theorem 9.8.2): *If μ_n , $n = 1, 2, \dots$ are probability measures on \mathbb{R}^N whose characteristic functions converge for all x to some $g(x)$, where g is continuous at 0 along each coordinate axis, then $\mu_n \rightarrow \mu$ weakly* to a probability measure μ with characteristic function g .*

We will also use Skorokhod's representation theorem:

Theorem 2.5. (cf. [Bi], Chapter 5)

Let $\mu_n, n \in \mathbb{N}$ be a sequence of probability measures on a metric space S such that μ_n converges weakly to some probability measure μ as $n \rightarrow \infty$. Suppose also that the support of μ is separable. Then there exist random variables Y_n, Y defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that the law of Y_n is μ_n , the law of Y is μ and such that $Y_n(\omega)$ converges to $Y(\omega)$ for all $\omega \in \Omega$.*

3. COMPLETENESS IN A FOURIER-BASED METRIC

A significant part of the paper [CT] is devoted to the study of a metric based on Fourier transform, defined on suitable spaces of probability measures:

Notation 3.1. ([CT], page 88) Fix a real number $s > 0$. For any pair of probability measures μ, ν on \mathbb{R}^N we let

$$d_s(\mu, \nu) = \sup_{x \in \mathbb{R}^N \setminus \{0\}} \frac{\hat{\mu}(x) - \hat{\nu}(x)}{|x|^s},$$

where $\hat{\mu}(x)$ is the Fourier transform (characteristic function) of μ .

In [CT], finiteness of the above expression for certain pairs of probability measures was proved:

Proposition 3.2. ([CT], Proposition 2.6): *Let $s > 0$ be given and let \mathcal{P}_s denote the space of all probability measures on \mathbb{R}^N with finite moments up to order $[s]$. The expression $d_s(\mu, \nu)$ is finite if μ, ν have equal moments up to order $[s]$ if $s \notin \mathbb{N}$ or up to order $s - 1$ if $s \in \mathbb{N}$.*

Remark 3.3. Note that the proof works for probability measures satisfying a weaker assumption, that is, those with characteristic functions differentiable up to order $[s]$ at 0. Indeed, from Taylor's formula with Peano remainder it can be easily seen that the expression $d_s(\mu, \nu)$ is finite if μ, ν have equal derivatives up to order $[s]$ at 0 if $s \notin \mathbb{N}$ or up to order $s - 1$ if $s \in \mathbb{N}$. More precisely, for $s \notin \mathbb{N}$ and probability measures $\mu, \nu \in \mathcal{P}_s$ with equal derivatives up to order $[s]$ at 0 we get that $\mu(x) - \nu(x)$ is of order $o(\|x\|^{[s]})$ in a neighborhood of 0, hence $d_s(\mu, \nu) < +\infty$. For $s \in \mathbb{N}$ and any two μ, ν with equal derivatives up to order $s - 1$, the expression $d_s(\mu, \nu)$ is still finite, but now the bound in a neighborhood of 0 involves also the difference between the derivative maps at 0 of order s of $\hat{\mu}$ and $\hat{\nu}$. In view of results of [Ro87], it would be interesting to find other possible conditions sufficient for finiteness of d_s .

Some subspaces of \mathcal{P}_s are complete with respect to the metric d_s . An example is presented in [CT]: given $s, a > 0$, let us denote by $X_{s,a,M}$ the set of probability measures $\mu \in \mathcal{P}_{s+a}(\mathbb{R}^N)$ such that $\int_{\mathbb{R}^N} v^\beta d\mu(v) = M_\beta \in \mathbb{R}_+$ for all multi-indices $|\beta| \leq [s]$ with M_β fixed numbers and $\int_{\mathbb{R}^N} v^{s+a} d\mu(v) \leq M_{s+a} \in \mathbb{R}_+$, where the set of all M_β and M_{s+a} is denoted simply by M .

Proposition 3.4. ([CT], Proposition 2.7): *The set $X_{s,a,M}$ endowed with the distance d_s is a complete metric space.*

As noticed in Remark 2.8 of [CT], the proof of the Proposition 3.4 given in that paper does not establish the completeness of the set of probability measures $\mu \in \mathcal{P}_m(\mathbb{R}^N)$ with $m \in \mathbb{N}$, such that $\int_{\mathbb{R}^N} v^\beta d\mu(v) = M_\beta \in \mathbb{R}_+$ for all multi-indices $|\beta| \leq m$ with M_β given,

endowed with the distance d_m . Therefore an open problem was formulated within the same remark as follows: “It would be nice to prove or rather disprove such statement at least for the d_2 distance.” The assumptions on μ_n are not enough to conclude anything about uniform behavior of their derivatives of order s . Nevertheless, below we are going to solve this problem by proving the following statement:

Theorem 3.5. *Let $s \in \mathbb{N}$ be fixed and let \mathcal{P}_s^- denote the space of all probability measures in \mathcal{P}_s which have equal moments up to order s . The metric space (\mathcal{P}_s^-, d_s) is complete.*

Proof. Let the sequence of measures $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathcal{P}_s^-$ be a Cauchy sequence with respect to the metric d_s . That is, for every $\varepsilon > 0$ there is an $n_\varepsilon \in \mathbb{N}$ such that for every $n, m > n_\varepsilon$ one has the inequality $\sup_{x \in \mathbb{R}^N \setminus 0} \frac{|\hat{\mu}_n(x) - \hat{\mu}_m(x)|}{|x|^s} < \varepsilon$.

Step 1. We will first show that μ_n converges weakly* to a probability measure $\mu \in \mathcal{P}_s^-$. To do this, observe (similarly to Proposition 3.4 in [CT]) that for any fixed $x \in \mathbb{R}^N$ the sequence $\{\hat{\mu}_n(x)\}_{n \in \mathbb{N}}$ is a Cauchy sequence in \mathbb{C} (for $x = 0$ the sequence is constant, of value 1). Hence the sequence of Fourier transforms $\{\hat{\mu}_n\}$ converges pointwise to some function g on \mathbb{R}^N with $g(0) = 1$. The equality of all moments of orders up to 1 for all μ_n imply that the functions $\hat{\mu}_n$ are uniformly Lipschitz on a neighborhood U of 0 in \mathbb{R}^N . By Ascoli’s theorem (cf. [Sch81b], Chapter VII, Section 6, Theorem 48, Corollary 1 and preceding examples), $\hat{\mu}_n$ converge uniformly in U to a continuous function g . By Lévy continuity theorem g is a Fourier transform of a probability measure μ and $\mu_n \rightarrow \mu$ weakly* as $n \rightarrow \infty$.

Step 2. We will now show that $d_s(\mu_n, \mu) \rightarrow 0$. Note that our assumptions imply that the sequence of complex-valued functions $f_n(x) = \frac{|\hat{\mu}_n(x) - \hat{\mu}(x)|}{|x|^s}$ is a Cauchy sequence in the space $\mathcal{C}^b(\mathbb{R}^N \setminus 0)$ of continuous bounded (complex-valued) functions in $\mathbb{R}^N \setminus 0$ endowed with the supremum norm. This is a complete metric space, so there exists a function $f \in \mathcal{C}^b(\mathbb{R}^N \setminus 0)$ such that f_n converge to f uniformly in $\mathbb{R}^N \setminus 0$. Hence for every $x \neq 0$ we have $|\hat{\mu}_n(x) - \hat{\mu}(x)| \rightarrow |x|^s f(x)$ as $n \rightarrow \infty$. The convergence of $\hat{\mu}_n$ to $\hat{\mu}$ implies that for every $x \neq 0$ one has $|x|^s f(x) = 0$, and so $f(x) = 0$ for every $x \neq 0$. This proves the claim that $d_s(\mu_n, \mu) \rightarrow 0$ as $n \rightarrow \infty$.

Step 3. Now we show that the limit measure μ has all the moments of order up to s . Take the random vectors Y_n, Y as in Skorokhod representation theorem. Let β be a multi-index with $|\beta| \leq s$. The moment $\int x_1^{\beta_1} \dots x_N^{\beta_N} d\mu_n$ is equal to $\int Y_{n,1}^{\beta_1} \dots Y_{n,N}^{\beta_N} d\mathbf{P}$. Since $Y_n \rightarrow Y$ pointwise as $n \rightarrow \infty$, we also have $Y_{n,1}^{\beta_1} \dots Y_{n,N}^{\beta_N} \rightarrow Y_1^{\beta_1} \dots Y_N^{\beta_N}$ as $n \rightarrow \infty$. By Fatou's lemma, $\int x_1^{\beta_1} \dots x_N^{\beta_N} d\mu = \int Y^\beta d\mathbf{P} \leq \liminf_{n \rightarrow \infty} \int Y_n^\beta d\mathbf{P} = M_\beta$, so μ has finite moments up to order s .

Step 4. It remains to show that all the moments of μ are equal to the corresponding moments of the measures μ_n . Note that the existence of all moments of order up to s implies that $\hat{\mu}$ is s times differentiable. Let $L_0 = 1$, $L_k = D^{(k)}\hat{\mu}(0)$, $k = 1, \dots, s$. From the assumption of equality of moments of corresponding orders for all μ_n , the mapping L_k is the same k -linear continuous symmetric mapping for all $n \in \mathbb{N}$. We already know that $\hat{\mu}(0) = 1$. Let $\varepsilon > 0$ be given. We estimate

$$\frac{\|\hat{\mu}(x) - L_0 + L_1.x + \dots + \frac{L_s}{s!} \cdot (x, \dots, x)\|}{\|x\|^s}$$

from above by

$$\frac{\|\hat{\mu}_n(x) - L_0 + L_1.x + \dots + \frac{L_s}{s!} \cdot (x, \dots, x)\|}{\|x\|^s} + \frac{\|\hat{\mu}_n(x) - \hat{\mu}(x)\|}{\|x\|^s},$$

where n is such that $\frac{\|\hat{\mu}_n(x) - \hat{\mu}(x)\|}{\|x\|^s} < \varepsilon/2$ (any such n will do). By Taylor's formula we can pick a $\delta > 0$ such that $\frac{\|\hat{\mu}_n(x) - L_0 + L_1.x + \dots + \frac{L_s}{s!} \cdot (x, \dots, x)\|}{\|x\|^s} < \varepsilon/2$ as $\|x\| < \delta$. This proves that $\|\hat{\mu}(x) - L_0 + L_1.x + \dots + \frac{L_s}{s!} \cdot (x, \dots, x)\| = o(\|x\|^s)$. By the converse to Taylor's theorem, $L_k = D^{(k)}\hat{\mu}(0)$ for all $k = 1, \dots, s$. This implies the equality of all partial derivatives of $\hat{\mu}$ at 0 with the derivatives of corresponding orders of $\hat{\mu}_n$ at 0, and hence of all the moments of corresponding orders up to s . □

Acknowledgment: I thank the anonymous referee for pointing out shortcomings of the preliminary version of this note.

REFERENCES

- [Bi] P. Billingsley: *Convergence of probability measures*, Second edition. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999. x+277 pp. ISBN: 0-471-19745-9 MR1700749

- [CT] Carrillo, J. A.; Toscani, G. Contractive probability metrics and asymptotic behavior of dissipative kinetic equations. *Riv. Mat. Univ. Parma* (7) 6 (2007), 75-198. MR2355628
- [Ca41] Cartan, Henri Sur les fondements de la théorie du potentiel. *Bull. Soc. Math. France* 69, (1941), 71-96. MR0015623
- [Ca45] Cartan, Henri Théorie du potentiel newtonien: énergie, capacité, suites de potentiels. *Bull. Soc. Math. France* 73, (1945), 74-106. MR0015622
- [CKL] U. Cegrell, S. Kołodziej, N. Levenberg: Two problems on potential theory for unbounded sets, *Math. Scand.* 83 (1998), no. 2, 265-276. MR1673930
- [Cho15] Y.-K. Cho: Absolute moments and Fourier-based probability metrics, preprint (2015), <http://arxiv.org/pdf/1510.08667.pdf>
- [De50] Deny, Jacques Les potentiels d'énergie finie. *Acta Math.* 82, (1950). 107-183. MR0036371
- [De51] Deny, Jacques Sur la définition de l'énergie en théorie du potentiel. *Ann. Inst. Fourier Grenoble* 2 (1950), 83-99 (1951). MR0044679
- [Du02] Dudley, R. M. Real analysis and probability. Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002. x+555 pp. ISBN: 0-521-00754-2 MR1932358
- [GSu] A. L. Gibbs and F. Edward Su, On choosing and bounding probability metrics, <http://arxiv.org/pdf/math/0209021.pdf> (2002)
- [GTW] Gabetta, E.; Toscani, G.; Wennberg, B.: Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation. *J. Statist. Phys.* 81 (1995), no. 5-6, 901-934. MR1361302
- [Ro87] L. V. Rozovskii: The accuracy of approximation of characteristic functions by polynomials, *J. Math. Sci.* 36:4 (1987), 532-535 MR0788196
- [Sch81a] L. Schwartz: Cours d'analyse. 1. [Course in analysis. 1] Second edition. Hermann, Paris, 1981. xxix+830 pp. ISBN: 2-7056-5764-9 MR0756814
- [Sch81b] L. Schwartz: Cours d'analyse. 2. [Course in analysis. 2] Second edition. Hermann, Paris, 1981. xxiii+475+21+75 pp. ISBN: 2-7056-5765-7 MR0756815
- [TV] Toscani, G.; Villani, C. Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *J. Statist. Phys.* 94 (1999), no. 3-4, 619-637. MR1675367
- [V] Villani, C. Optimal transport, old and new. Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer, Berlin, 2009. MR2459454

MATHEMATICAL REVIEWS, 416 FOURTH ST., ANN ARBOR, MI 48103, USA
E-mail address: stawiska@umich.edu